



RECEIVED  
JAN 31 2002  
TECH CENTER 1600/2900

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application of: S. Stubbs, et al. Group Art Unit: To be assigned  
Serial Number: 09/967,301 Examiner: To be assigned  
Filing Date: September 28, 2001  
Title: Fluorescent Proteins

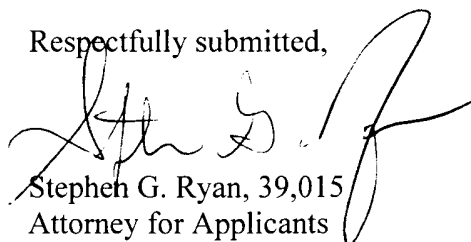
COMPLETION OF CLAIM FOR PRIORITY

Assistant Commissioner for Patents  
Washington, D.C. 20231

Sir:

Applicants hereby submit the official certified copy of the priority document number **GB 0109858.1** in connection with the above identified application, benefit of which is claimed in the declaration of this application. The Examiner is most respectfully requested to acknowledge receipt of this certified copy in the next Official Office Action.

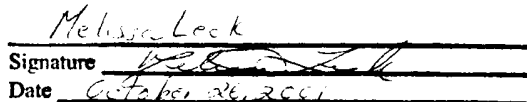
Respectfully submitted,

  
Stephen G. Ryan, 39,015  
Attorney for Applicants

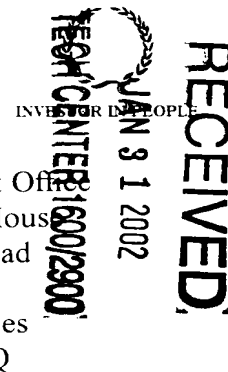
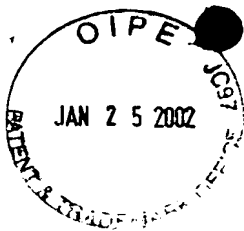
Amersham Pharmacia Biotech, Inc.  
800 Centennial Avenue  
P. O. Box 1327  
Piscataway, New Jersey 08855-1327

Tel: (732) 457-8071  
Fax: (732) 457-8463

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner of Patents and Trademarks, Washington, D.C. 20231, on 10-26-01.

  
Signature \_\_\_\_\_  
Date October 26, 2001

\_\_\_\_\_



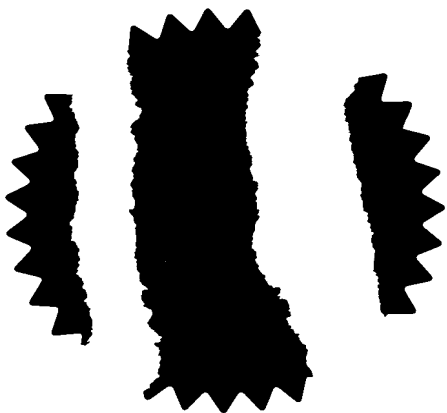
The Patent Office  
Concept House  
Cardiff Road  
Newport  
South Wales  
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.



Signed

Dated 16 October 2001



## Patents Form 1/77

Patents Act 1977  
(Rule 16)

23 APR 2001

2345678 5633719-1 003901  
PA1/7700 0.01-0109858.1

## Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road  
Newport  
South Wales  
NP10 8QQ

1. Your reference PA0111

2. Patent application number  
0109858.1

23 APR 2001

3. Name and postcode of the or of each applicant (underline all surnames)

AMERSHAM PHARMACIA BIOTECH UK LIMITED  
Amersham Place  
Little Chalfont  
Buckinghamshire  
HP7 9NA

Patents ADP number (if you know it)

If the applicant is a corporate body, give the country/state of its incorporation

GB

7427206000

4. Title of the invention

FLUORESCENT PROTEINS

5. Name of your agent (if you have one)

"Address for service" in the United Kingdom to which all correspondence should be sent (including the postcode)

HAMMER, Catriona, MacLeod;  
ROLLINS, Anthony, John and  
HAMMETT, Audrey, Grace, Campbell  
Nycomed Amersham plc  
Amersham Laboratories  
White Lion Road  
Amersham  
Buckinghamshire HP7 9LL

Patents ADP number (if you know it)

743303000

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (if you know it) the or each application number

Country

Priority application number  
(if you know it)Date of filing  
(day / month / year)

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application

Date of filing  
(day / month / year)

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? (Answer 'Yes' if:

YES

- a) any applicant named in part 3 is not an inventor, or
  - b) there is an inventor who is not named as an applicant, or
  - c) any named applicant is a corporate body.
- See note (d))

Patents Form 1/77

0004751 23-Apr-01 02:06

## Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form

Description 23

Claim(s) 5

Abstract

Drawing(s) 7 -

Sequence drawings not drawings 1 Sci A

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (Patents Form 7/77)

Request for preliminary examination and search (Patents Form 9/77) 1

Request for substantive examination (Patents Form 10/77)

Any other documents (please specify)

11.

I/We request the grant of a patent on the basis of this application.

Signature   
HAMMER, Catriona, MacLeodDate  
23 April 2001

12. Name and daytime telephone number of person to contact in the United Kingdom HALLS, Jennie 01494 545064

**Warning**

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

**Notes**

- If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.
- Write your answers in capital letters using black ink or you may type them.
- If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- Once you have filled in the form you must remember to sign and date it.
- For details of the fee and ways to pay please contact the Patent Office.

Patents Form 1/77

0004751 23-Apr-01 02:06

DUPLICATE

FLUORESCENT PROTEINS

The present invention relates to novel variants of the fluorescent protein GFP having improved fluorescence properties.

5

The use of Green Fluorescent Protein (GFP) derived from *Aequorea victoria* has revolutionised research into many cellular and molecular-biological processes. GFP allows researchers to label proteins within cells with an intrinsic fluor, so obviating the requirement to perform chemical labelling of proteins, and  
10 allowing development of assays of biological processes in intact living cells.

US 5491084 describes the use of GFP as a biological reporter. Early applications of GFP as a biological reporter (Chalfie et al. Science, (1994), 263, 802-5; Chalfie, et al, Photochem.Photobiol., (1995), 62(4), 651-6) used wild type  
15 (native) GFP (wtGFP), but these studies quickly demonstrated two areas of deficiency of wtGFP as a reporter for use in mammalian cells. Firstly, the protein being derived from a poikilothermic marine organism does not undergo protein folding efficiently when expressed in mammalian cells cultured at 37 °C, resulting in weak fluorescence. Secondly, the spectral characteristics of the wtGFP are  
20 not ideally suited to use as a cellular reporter, requiring excitation with electromagnetic radiation in the near-UV range, which is potentially damaging to living cells.

Consequently, significant effort has been expended to produce variant  
25 mutated forms of GFP with properties more suitable for use as an intracellular reporter.

A number of mutated forms of GFP with altered spectral properties have been described. A variant-GFP (Heim et al. (1994) Proc.Natl.Acad.Sci. 91, 12501) contains a Y66H mutation which blue-shifts the excitation and emission  
30 spectrum of the protein. However, this protein is only weakly fluorescent and requires potentially damaging UV excitation.

-2-

A further mutant of GFP (Heim et al, Nature, (1995), 373, 663-664) contains a S65T mutation which red-shifts the optimum excitation and emission wavelengths relative to wtGFP and which is 4-6 fold brighter than wtGFP when expressed as a recombinant protein at 25 °C. However, this variant does not  
5 yield bright fluorescence when expressed in hosts cultured at 37 °C.

Ehrig et al (FEBS Lett., (1995), 367,163-6) describe two mutations of GFP, T203I and E222G, which individually delete one of the excitation maxima of wtGFP. The E222G mutation deletes the near-UV excitation peak at 395 nm and  
10 produces a red-shift in the excitation peak at 475 nm to 481 nm. The emission peak for this mutant protein is at 506 nm.

WO96/27675 describes two variant GFPs, obtained by random mutagenesis and subsequent selection for brightness, which contain the  
15 mutations V163A and V163A+S175G, respectively. These variants were shown to produce more efficient expression in plant cells relative to wtGFP and to increase the thermotolerance of protein folding. The double mutant V163A+S175G was observed to be brighter than the variant containing the single V163A mutant alone; however this mutant exhibits an undesirable blue-shifted  
20 excitation peak.

A further mutant, termed cycle-3, generated by molecular evolution through DNA shuffling (Crameri, A. et al, Nature Biotechnology, (1996), 14, 315-9) is available commercially from Invitrogen Inc. Cycle-3-GFP contains three  
25 mutations (F99S+M153T+V163A) which increase whole cell fluorescence approximately 42 fold when compared with wtGFP. However, this mutant retains the near-UV excitation maximum of the wtGFP, making it less suitable as a reporter for use in living cells.

30 The above mutations effectively address some of the spectral deficiencies of wtGFP as a biological reporter in providing variant forms of GFP which are compatible with lower energy excitation and which emit at wavelengths compatible with detection instrumentation commonly in use for measuring



-3-

biological reporters. However, such mutations do not address the problem of inefficient folding and chromophore formation when wtGFP or spectral variants are expressed in hosts requiring growth at temperatures significantly greater than ambient.

5

US 6172188 describes variant GFPs wherein the amino acid in position 1 preceding the chromophore has been mutated to provide an increase of fluorescence intensity. Such mutations include F64I, F64V, F64A, F64G and F64L, with F64L being the preferred mutation. These mutants result in a substantial increase in the intensity of fluorescence of GFP without shifting the excitation and emission maxima. F64L-GFP has been shown to yield an approximate 6-fold increase in fluorescence at 37 °C due to shorter chromophore maturation time.

15

In addition to the single mutants or randomly derived combinations of mutations described above, a variety of variant-GFPs have been created which contain two or more mutations deliberately selected from those described above and other mutations, and which seek to combine the advantageous properties of the individual mutations to produce a protein with expression and spectral properties which are suited to use as a sensitive biological reporter in mammalian cells.

25

One mutant, commonly termed EGFP, available commercially from Clontech Inc., contains the mutations F64L and S65T (Cormack, B.P. et al, Gene, (1996), 173, 33-38). These mutations when combined, confer an approximate 35-fold increase in brightness over wtGFP and the spectral characteristics permit excitation and detection of EGFP with commonly used fluorescein excitation (488 nm) and emission filters (505 nm-530 nm). EGFP has been optimised for expression in mammalian systems, having been constructed with preferred mammalian codons.

30

US 6194548 (Osumi et al) discloses GFPs with improved fluorescence and folding characteristics at 37 °C that contain, at least, the changes F64L and V163A and S175G.

5 US 6077707 (Tsien et al) describes a blue fluorescent protein (BFP) containing the F64L mutation in combination with Y66H and US 6194548 (Osumi et al) describes a further BFP containing the F64L, Y66H, Y145F and L236R substitutions.

10 The present invention provides novel engineered derivatives of green fluorescent protein (GFP) which exhibit enhanced fluorescence relative to wtGFP when expressed in non-homologous cells at temperatures above 30°C, and when excited at about 490 nm compared to the parent proteins, i.e. wtGFP. Mutant GFPs produced using the method of the invention provide a means for  
15 detecting GFP reporters in mammalian cells at lower levels of expression and/or increased sensitivity relative to wtGFP. This greatly improves the usefulness of fluorescent proteins in studying cellular functions in living cells. The multiply-mutated GFPs of this invention have fluorescence properties which are not predictable from the properties of the individual mutations when studied in  
20 isolation. Furthermore, the inventors have surprisingly found that certain GFPs according to the present invention, which do not contain any mutations in the chromophore region relative to wtGFP, exhibit enhanced fluorescence compared with mutant GFPs described previously.

25 In a first aspect of the invention, there is provided a fluorescent protein which is derived from Green Fluorescent Protein (GFP) or any functional GFP analogue and has an amino acid sequence which is modified by amino acid substitution compared with the amino acid sequence of wild type Green Fluorescent Protein said modified fluorescent protein comprising:

- 30 i) an amino acid substitution at position F64;  
ii) a single amino acid substitution at a position selected from the group consisting of positions S65 and E222; and

-5-

iii) a single amino acid substitution at a position selected from the group consisting of positions V163 and S175; wherein said modified GFP has a different excitation spectrum and/or emission spectrum compared with wild type GFP.

5

Suitably, the amino acid F at position 64 may be substituted by an amino acid selected from the group consisting of L, I, V, A and G, thereby providing F64L, F64I, F64V, F64A, or F64G substitutions. Preferably, the amino acid substitution at position 64 is the F64L substitution.

10

In embodiments where the amino acid S at position 65 is substituted, it is suitably substituted by an amino acid selected from the group consisting of G, A, L, C, V, I and T, thereby providing S65G, S65A, S65L, S65C, S65V, S65I or S65T substitutions. Preferably, the amino acid substitution at position 65 is the S65T substitution.

15

In embodiments where the amino acid E at position 222 is substituted, it is suitably substituted by an amino acid selected from the group consisting of G, A, V, L, I, F, S, T, N and Q, thereby providing E222G, E222A, E222V, E222L, E222I, E222F, E222S, E222T, E222N or E222Q substitutions. Preferably, the amino acid substitution at position 222 is E222G.

20

In one preferred embodiment of the first aspect, the amino acid V is substituted by A at position 163. In a second preferred embodiment of the first aspect, the amino acid S is substituted by G at position 175.

25

Suitably, the novel fluorescent proteins exhibit high fluorescence in cells expressing them when said cells are incubated at a temperature of 30 °C or above, preferably at a temperature of from 32 °C to 39 °C, more preferably from 35 °C to 38 °C and most preferably at a temperature of about 37 °C.

30

Preferably, the fluorescent protein according to the first aspect has an amino acid sequence which is modified by amino acid substitution compared with

-6-

the amino acid sequence of wild type Green Fluorescent Protein having the sequence: SEQ ID No.2.

5 A preferred protein according to the present invention is a protein in which, in relation to SEQ ID No.2 of GFP, the amino acid F at position 64 has been substituted by L, the amino acid V at position 163 has been substituted by A and the amino acid E at position 222 has been substituted by G, and is shown herein as having the amino acid sequence as set forth in SEQ ID No.3.

10 An alternative preferred protein according to the present invention is a protein in which, in relation to SEQ ID No.2 of GFP, the amino acid F at position 64 has been substituted by L, the amino acid S at position 175 has been substituted by G and the amino acid E at position 222 has been substituted by G, and is shown herein as having the amino acid sequence as set forth in SEQ ID  
15 No.4.

A further preferred protein according to the present invention is a protein in which, in relation to SEQ ID No.2 of GFP, the amino acid F at position 64 has been substituted by L, the amino acid S at position 65 has been substituted by T  
20 and the amino acid S at position 175 has been substituted by G, and is shown herein as having the amino acid sequence as set forth in SEQ ID No.5.

A further preferred protein according to the present invention is a protein in which, in relation to SEQ ID No.2 of GFP, the amino acid F at position 64 has  
25 been substituted by L, the amino acid S at position 65 has been substituted by T and the amino acid V at position 163 has been substituted by A, and is shown herein as having the amino acid sequence as set forth in SEQ ID No.6.

Suitably, the GFP or functional GFP-analogue used to derive the  
30 fluorescent protein may be obtained from any convenient source. For example, native GFP derived from species of the genus *Aequorea*, suitably *Aequorea victoria*. The chromophore in wtGFP from *Aequorea victoria* is at positions 65-67

-7-

of the predicted primary amino acid sequence (SEQ ID No.2). In a preferred embodiment, the GFP is derived from *Aequorea victoria*.

The modified proteins of the present invention may be produced by  
5 introducing mutations in a sequence of the nucleic acid that encodes the protein. As used herein, a preferred sequence of the gene encoding wtGFP is derived from *Aequorea victoria*, published by Chalfie et al, (Science, (1994), 263, 802-5) disclosed as SEQ. ID. No1 (Figure 1). The corresponding amino acid sequence is shown in SEQ. ID. No2 (Figure 2). Alternative sequences of the GFP gene  
10 may be used, for example, the nucleotide (and predicted amino acid) sequences of the GFP gene described by Prasher et al, (Gene (1992), 111, 229) and the sequences as disclosed in WO 97/11094. In addition, alternative gene sequences that encode the fluorescent protein may incorporate a consensus Kozak nucleotide sequence (Kozak, M., Cell (1986), 44, 283), or preferred  
15 mammalian codons, to provide improved translation in mammalian systems. The nucleotide sequence corresponding to the fluorescent protein may also encode useful restriction enzyme sites and additional elements such as target sites for enzymes and purification tags. Methods for incorporation of a Kozak region, preferred mammalian codons, restriction enzyme sites, enzyme sites and  
20 purification tags are well known in the art and may result in the incorporation of amino acid residues and a change in numbering of amino acid residues in the fluorescent protein relative to the wtGFP numbering in the sequence provided.

Herein, the abbreviations used for the amino acids are those stated in  
25 J.Biol.Chem., (1968), 243, 3558.

In a second aspect of the invention, there is provided a fusion compound comprising a protein of interest fused to a fluorescent protein which is derived from Green Fluorescent Protein (GFP) or any functional GFP analogue and has  
30 an amino acid sequence which is modified by amino acid substitution compared with the amino acid sequence of wild type Green Fluorescent Protein said modified fluorescent protein comprising:

- i) an amino acid substitution at position F64;

-8-

ii) a single amino acid substitution at a position selected from the group consisting of positions S65 and E222; and

iii) a single amino acid substitution at a position selected from the group consisting of positions V163 and S175;

5 wherein said modified GFP has a different excitation spectrum and/or emission spectrum compared with wild type GFP.

In the context of the present invention, the term "protein of interest" is intended also to encompass polypeptides and peptide fragments. Examples of  
10 such proteins of interest may be selected from the group consisting of: NF $\kappa$ B and subunits thereof, RAC1, PLC domains, MAPKAP2, PKC, Cytochrome C, RHO,  $\beta$ -actin, STAT6, protein kinase C isotypes, LAMP1/2 TGN, ATP7A TGN and GLUT4.

15 In a third aspect of the present invention there is provided a nucleic acid molecule comprising a nucleotide sequence encoding a fluorescent protein which is derived from Green Fluorescent Protein (GFP) or any functional GFP analogue and has an amino acid sequence which is modified by amino acid substitution compared with the amino acid sequence of wild type Green Fluorescent Protein  
20 said modified fluorescent protein comprising:

i) an amino acid substitution at position F64;

ii) a single amino acid substitution at a position selected from the group consisting of positions S65 and E222; and

25 iii) a single amino acid substitution at a position selected from the group consisting of positions V163 and S175;

wherein said modified GFP has a different excitation spectrum and/or emission spectrum compared with wild type GFP.

Preferably, the nucleic acid molecule according to the third aspect  
30 encodes a fluorescent protein having an amino acid sequence which is modified by amino acid substitution compared with the amino acid sequence of wild type Green Fluorescent Protein having the sequence: SEQ ID No.2.

-9-

In a particular embodiment of the third aspect, the nucleic acid molecule comprises a nucleotide sequence encoding a fluorescent protein derived from Green Fluorescent Protein (GFP) or any functional GFP analogue according to the invention fused to a nucleotide sequence encoding a protein of interest.

5

Preferably, the nucleic acid molecule is a construct comprising a DNA sequence.

Preferably, the nucleic acid molecule encodes a fluorescent protein having  
10 an amino acid sequence selected from the group consisting of SEQ ID No.3, SEQ ID No.4, SEQ ID No.5 and SEQ ID No.6.

As is well known, a single amino acid may be encoded by more than one nucleotide codon and thus each of the above nucleotide sequences may be  
15 modified to produce an alternative nucleotide sequence that encodes the same peptide. Thus, the preferred embodiments of the invention include alternative DNA sequences that encode the preferred proteins as previously described. It is to be understood that the preferred proteins (and the nucleic acid sequences from which they are derived), may include additional residues, particularly N- and C-  
20 terminal amino acids, or 5'- or 3'-nucleotide sequences, and still be essentially as described herein.

Suitably, the DNA construct encoding the novel fluorescent proteins may be prepared synthetically by established methods, e.g. the phosphoramidite  
25 method described by Beaucage and Caruthers, (Tetrahedron Letters (1981), 22, 1859-1869), or the method described by Matthes et al., (EMBO Journal (1984), 3, 801-805). According to the phosphoramidite method, oligonucleotides are synthesized, e.g. in an automatic DNA synthesizer, purified, annealed, ligated and cloned into suitable vectors.

30

The DNA construct encoding the fluorescent protein may also be prepared by recombinant DNA methodology, for example cDNA cloning. See for example,

-10-

Sambrook, J. et al (1989) Molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory Press.

The DNA construct may also be prepared by polymerase chain reaction  
5 (PCR) using specific primers, for instance as described in US 4683202 or by  
Saiki et al (Science (1988), 239, 487-491). A recent review of PCR methods may  
be found in PCR Protocols, (1990), Academic Press, San Diego, California, USA.

The gene sequence encoding the fluorescent protein may be joined in-  
10 frame with a gene encoding the protein of interest and the desired fusion protein  
produced when inserted into an appropriate expression vector. For example,  
polymerase chain reaction or complementary oligonucleotides may be employed  
to engineer a polynucleotide sequence corresponding to the fluorescent protein,  
5' or 3' to the gene sequence corresponding to the protein of interest.  
15 Alternatively, the same techniques may be used to engineer a polynucleotide  
sequence corresponding to the fluorescent protein sequence 5' or 3' to the  
multiple cloning site of an expression vector prior to insertion of a gene sequence  
encoding the protein of interest. The polynucleotide sequence corresponding to  
the fluorescent protein sequence may comprise additional nucleotide sequences  
20 to include cloning sites, linkers, transcription and translation initiation and/or  
termination signals, labelling and purification tags.

In a fourth aspect, there is provided an expression vector comprising  
suitable expression control sequences operably linked to a nucleic acid molecule  
25 according to the present invention. The DNA construct of the invention may be  
inserted into a recombinant vector, which may be any vector that may  
conveniently be subjected to recombinant DNA procedures. The choice of vector  
will often depend on the host cell into which it is to be introduced. Thus, the  
vector may be an autonomously replicating vector, ie. a vector which exists as an  
30 extrachromosomal entity, the replication of which is independent of chromosomal  
replication, eg. a plasmid. Alternatively, the vector may be one which, when  
introduced into a host cell, is integrated into the host cell genome and replicated  
together with the chromosome(s) into which it has been integrated.



-11-

The vector is preferably an expression vector in which the DNA sequence encoding a fluorescent protein of the invention is operably linked to additional segments required for transcription of the DNA. In general, the expression vector  
5 is derived from plasmid or viral DNA, or may contain elements of both. The term, "operably linked" indicates that the segments are arranged so that they function in concert for their intended purposes, e.g. transcription initiates in a promoter and proceeds through the DNA sequence coding for the fluorescent protein of the invention.

10

The promoter may be any DNA sequence which shows transcriptional activity in a suitable host cell of choice, (eg. a bacterial cell, a mammalian cell, a yeast cell, or an insect cell) for expressing a fluorescent protein. The promoter  
15 may be derived from genes encoding proteins either homologous or heterologous to the host cell.

Examples of suitable promoters for directing the transcription of the DNA sequence encoding the fluorescent protein of the invention in mammalian cells are the CMV promoter (US 5168062, US5385839), Ubiquitin C promoter (Wulff,  
20 M. et al., FEBS Lett. (1990), 261, 101-105), SV40 promoter (Subramani et al., Mol. Cell Biol. (1981), 1, 854-864) and MT-1 (metallothionein gene) promoter (Palmiter et al., Science (1983), 222, 809-814). An example of a suitable promoter for use in insect cells is the polyhedrin promoter (US 4745051; Vasuvedan et al., FEBS Lett., (1992) 311, 7-11). Examples of suitable promoters  
25 for use in yeast host cells include promoters from yeast glycolytic genes (Hitzeman et al., J. Biol. Chem., (1980), 255, 12073-12080; Alber and Kawasaki, J. Mol. Appl. Gen., (1982), 1, 419-434) or alcohol dehydrogenase genes (Young et al., in Genetic Engineering of Microorganisms for Chemicals (Hollaender et al, eds.), Plenum Press, New York, 1982), or the TPI1 (US 4599311) or ADH2-4c  
30 (Russell et al., Nature, (1983), 304, 652-654) promoters.

Examples of suitable promoters for use in bacterial host cells include the promoter of the *Bacillus stearothermophilus* maltogenic amylase gene, the

-12-

*Bacillus licheniformis* alpha-amylase gene, the *Bacillus amyloliquefaciens* BAN amylase gene, the *Bacillus subtilis* alkaline protease gene, or the *Bacillus pumilus* xylosidase gene, or the phage Lambda PR or PL promoters or the *Escherichia coli* *lac*, *trp* or *tac* promoters.

5

The DNA sequence encoding the novel fluorescent proteins of the invention may also, if necessary, be operably connected to a suitable terminator, such as the human growth hormone terminator (Palmiter et al., op. cit.) or (for fungal hosts) the TPI1 (Alber and Kawasaki, op. cit.) or ADH3 (McKnight et al., op. cit.) terminators. The vector may further comprise elements such as  
10 polyadenylation signals (e.g. from SV40 or the adenovirus 5 Elb region), transcriptional enhancer sequences (e.g. the SV40 enhancer) and translational enhancer sequences (e.g. the ones encoding adenovirus VA RNAs).

15

The vector may further comprise a DNA sequence enabling internal ribosomal entry and expression of two proteins from one bicistronic transcript mRNA molecule. For example, the internal ribosomal entry sequence from the encephalomyocarditis virus (Rees S, et al, BioTechniques (1996), 20, 102-110 and US 4937190).

20

The recombinant vector may further comprise a DNA sequence enabling the vector to replicate in the host cell in question. An example of such a sequence (when the host cell is a mammalian cell) is the SV40 origin of replication.

25

When the host cell is a yeast cell, examples of suitable sequences enabling the vector to replicate are the yeast plasmid 2 $\mu$  replication genes REP 1-3 and origin of replication.

30

The vector may also comprise selectable markers, such as a gene that confers resistance to a drug, e.g. ampicillin, kanamycin, tetracyclin, chloramphenicol, puromycin, neomycin or hygromycin.

-13-

The procedures used to ligate the DNA sequences coding for the fluorescent protein of the invention, the promoter and optionally the terminator and/ or targeting sequence, respectively, and to insert them into suitable vectors containing the information necessary for replication, are well known to persons skilled in the art (eg. Sambrook et al., op.cit.).

In a fifth aspect of the invention, there is provided a host cell transformed or transfected with a DNA construct comprising an expression vector according to the present invention.

The DNA construct or the recombinant vector of the invention is suitably introduced into a host cell which may be any cell which is capable of expressing the present DNA construct and includes bacteria, yeast and higher eukaryotic cells (Unger, T.F., The Scientist (1997), 11(17), 20-23; Smith, C., The Scientist (1998), 12(22): 20; Smith, C., The Scientist (1998), 12(3), 18; Fernandez, J.M. & Hoeffler, J.P., Gene Expression Systems- using nature for the art of expression, Academic Press 1999).

Examples of bacterial host cells which, on cultivation, are capable of expressing the DNA construct of the invention are Gram-positive bacteria, eg. species of *Bacillus* or Gram-negative bacteria such as *E. coli*. The transformation of the bacteria may be effected by using competent cells in a manner known per se (cf. Sambrook et al., supra).

Examples of suitable mammalian cell lines are the HEK293 and the HeLa cell lines, primary cells, and the COS (e.g. ATCC CRL 1650), BHK (eg. ATCC CRL 1632, ATCC CCL 10), CHL (e.g. ATCC CCL39) or CHO (eg. ATCC CCL 61) cell lines. Methods of transfecting mammalian cells and expressing DNA sequences introduced in the cells are described in eg. Kaufman and Sharp, J. Mol. Biol., (1982), 159, 601-621; Southern and Berg, J. Mol. Appl. Genet., (1982), 1, 327-341; Loyer et al., Proc. Natl. Acad. Sci. USA, (1982), 79, 422-426; Wigler et al., Cell, (1978), 14, 725; Corsaro and Pearson, Somatic Cell Genetics, (1981),

-14-

Z, 603, Graham and van der Eb, Virology (1973), 52, 456; and Neumann et al., EMBO J., (1982), 1, 841-845.

Examples of suitable yeast cells include cells of *Saccharomyces* spp. or  
5 *Schizosaccharomyces* spp., in particular strains of *Saccharomyces cerevisiae* or  
*Saccharomyces kluyveri*. Methods for transforming yeast cells with heterologous  
DNA and producing heterologous polypeptides therefrom are described, e.g. in  
US 4599311, US 4931373, US 4870008, US 5037743, and US 4845075, all of  
which are hereby incorporated by reference. Transformed cells are selected by a  
10 phenotype determined by a selectable marker, commonly drug resistance or the  
ability to grow in the absence of a particular nutrient, e.g. leucine. A preferred  
vector for use in yeast is the POT1 vector disclosed in US 4931373. The DNA  
sequence encoding the fluorescent protein of the invention may be preceded by a  
signal sequence and optionally a leader sequence, e.g. as described above.  
15 Further examples of suitable yeast cells are strains of *Kluyveromyces*, such as *K.*  
*lactis*, *Hansenula*, e.g. *H. polymorpha*, or *Pichia*, e.g. *P. pastoris* (cf. Gleeson et  
al., J. Gen. Microbiol., (1986), 132, 3459-3465; US 4882279).

Transformation of insect cells and production of heterologous polypeptides  
20 therein may be performed as described in US 4745051; US 4879236; US  
5155037; US 5162222; EP 397485, all of which are incorporated herein by  
reference. The insect cell line used as the host may suitably be a *Lepidoptera*  
cell line, such as *Spodoptera frugiperda* cells or *Trichoplusia ni* cells (cf. US  
5077214). Culture conditions may suitably be as described in, for instance, WO  
25 89/01029 or WO 89/01028, or any of the aforementioned references.

In a sixth aspect, the invention provides a method for preparing a Green  
Fluorescent Protein (GFP) or a functional GFP analogue according to the present  
invention, the method comprising cultivating a host cell transformed or  
30 transfected with a nucleotide sequence according to the invention and obtaining  
therefrom the polypeptide expressed by said nucleotide sequence.

-15-

Suitably, the transformed or transfected host cells as described above are cultured in a suitable nutrient medium under conditions permitting the expression of a DNA construct according to the invention, after which the cells may be used in the screening method of the invention. Alternatively, the cells may be  
5 disrupted after which cell extracts and/or supernatants may be analysed for fluorescence and/ or used to purify the GFP or functional GFP analogue of the invention.

The medium used to culture the cells may be any conventional medium  
10 suitable for growing the host cells, such as minimal or complex media containing appropriate supplements. Suitable media are available from commercial suppliers or may be prepared according to published protocols (eg. in catalogues of the American Type Culture Collection; Sambrook et al., supra).

For example, a fusion protein comprising glutathione S-transferase (GST)  
15 and GFP can be constructed and expressed in *E. coli*. The GFP may be joined in-frame to the C-terminus of GST in a pGEX plasmid vector (Amersham Pharmacia Biotech). Recombinant production of the fusion protein is carried out utilising a standard *E. coli* expression host, followed by purification employing  
20 glutathione affinity chromatography and removal of the GST tag by proteolytic cleavage.

In a seventh aspect of the present invention, there is provided a method of measuring the expression of a protein of interest in a cell. The method  
25 comprises: i) introducing into a cell a nucleic acid molecule comprising a nucleotide sequence encoding a fluorescent protein which is derived from Green Fluorescent Protein (GFP) or any functional GFP analogue according to the present invention said nucleic acid molecule being operably linked to and under the control of an expression control sequence which moderates expression of  
30 said protein of interest; ii) culturing the cell under conditions suitable for the expression of the protein of interest; and iii) detecting the fluorescence emission of the Green Fluorescent Protein (GFP) or a functional GFP analogue as a means of measuring the expression of the protein of interest.

In an eighth aspect of the present invention, there is provided a method of determining the cellular and/or extracellular localisation of a protein of interest which method comprises:

- 5 i) introducing into a cell a nucleic acid molecule comprising a nucleotide sequence encoding a Green Fluorescent Protein (GFP) or a functional GFP analogue according to the invention fused to a nucleotide sequence encoding a protein of interest, said nucleic acid molecule being operably linked to and under the control of a suitable expression control sequence;
- 10 ii) culturing said cell under conditions suitable for the expression of said protein of interest; and
- iii) determining the cellular and/or extracellular localisation of said protein of interest by detecting the fluorescence emission by optical means.

15 The fluorescent proteins of the present invention may also be used in a method to detect and compare the effect of a test substance on the regulation of expression and/or translocation of two or more different proteins of interest in a cell. Alternatively, they may be used in a method to compare the expression of a protein of interest and the simultaneous activity of an expression control

20 sequence in response to a test substance. The fluorescent proteins may also be used in a method to compare the activity of two or more expression control sequences in a cell in response to a test substance. Such methods may be performed in the presence and in the absence of a test substance whose effect on the process is to be measured. For example, one detectable reporter

25 molecule may be used as an internal reference and another as a variable marker, since regulated expression of a gene can be monitored quantitatively by fusion of an expression control sequence to a DNA construct encoding, eg. F64L-S175G-E222G-GFP, measuring the fluorescence, and normalising it to the fluorescence of a constitutively expressed spectrally distinct fluorescent molecule. The

30 constitutively expressed spectrally distinct fluorescent molecule, for example BFP, acts as an internal reference.

-17-

Thus, in a ninth aspect of the present invention, there is provided a method of comparing the effect of one or more test substance(s) on the expression and/or localisation of one or more different protein(s) of interest in a cell which method comprises:

5

i) introducing into a cell:

a) a nucleic acid molecule comprising a nucleotide sequence encoding a Green Fluorescent Protein (GFP) or a functional GFP analogue according to the invention optionally fused to a nucleotide sequence encoding a first protein of interest, said nucleic acid molecule being operably linked to and under the control of a first expression control sequence; and optionally,

b) at least one different nucleic acid molecule encoding a protein reporter molecule optionally fused to a different protein of interest, each said nucleic acid molecule being operably linked to and under the control of a second expression control sequence wherein said protein reporter molecule has or is capable of generating an emission signal which is spectrally distinct from that of said Green Fluorescent Protein (GFP) or a functional GFP analogue;

ii) culturing said cells under conditions suitable for the expression of said protein(s) of interest in the presence and absence of said test substance(s);

iii) determining the expression and/or localisation of said protein(s) of interest in said cells by detecting the fluorescence emission by optical means; and

iv) comparing the fluorescence emission obtained in the presence and absence of said test substance(s) to determine the effect of said test substance(s) on the expression and/or localisation of said protein(s) of interest.

25

In a preferred embodiment of the ninth aspect, samples of said cells in a fluid medium are introduced into separate vessels for each of said test substances to be studied.

Preferably, the first and second expression control sequences are different.

30

-18-

Suitably, the protein reporter molecule may be selected from the group consisting of fluorescent proteins and enzymes. Preferred fluorescent proteins are those which have a spectrally distinguishable emission wavelength compared with the emission wavelength of the fluorescent proteins according to the present invention, for example, BFP. Suitable enzyme reporters are those which are suitable for generating a detectable (eg. a luminescent or fluorescent) signal in a substrate. Suitable enzyme/substrates include: luciferase/luciferin;  $\beta$ -galactosidase/DDAO galactoside;  $\beta$ -galactosidase/fluorescein di- $\beta$ -D-galactopyranoside; alkaline phosphatase/Atttophos.

10

In the methods of the invention, the fluorescence of cells transformed or transfected with the DNA construct according to the invention may suitably be measured by optical means in for example; a spectrophotometer, a fluorimeter, a fluorescence microscope, a cooled charge-coupled device (CCD) imager (such as a scanning imager or an area imager), a fluorescence activated cell sorter, a confocal microscope or a scanning confocal device, where the spectral properties of the cells in culture may be determined as scans of light excitation and emission.

20

The fluorescent proteins of the present invention have many additional applications, for example:

i) Use as a non-toxic marker for selection of transfected cells containing an expression vector encoding at least the fluorescent protein of the invention. The fluorescent emission may be used to isolate transfected cells thereby overcoming the need for selection with toxic molecules such as antibiotics.

25

ii) Use as a protein label in living and fixed cells. The novel proteins exhibit strong fluorescence and are a suitable label for proteins present at low concentrations. Since no substrate is needed and visualization of the fluorescent protein does not damage the cells, dynamic analysis can be performed.

30



-19-

iii) Use as marker in cell or organelle fusion. By labelling one or more cells or organelles with the novel proteins, for example, F64L-S175G-E222G-GFP, and other cells or organelles with same or another fluor, fusions such as heterokaryon formation can be monitored.

5

iv) Translocation of proteins fused to the novel proteins of the invention can be visualised. The translocation of intracellular proteins to a specific organelle can be visualised by fusing the protein of interest to a fluorescent protein, for example, F64L-S175G-E222G-GFP and labelling the organelle with another  
10 fluorescent molecule, eg. fluorescent protein. Translocation can then be detected as a spectral shift in the fluorescent proteins in the specific organelle.

v) Use as a secretion marker. By fusion of a fluorescent protein of the invention to a signal peptide or a peptide to be secreted, secretion may be  
15 followed in living cells.

vi) Use as genetic reporter or protein tag in transgenic animals. Due to the strong fluorescence of the novel proteins, they are suitable as tags for proteins and gene expression, since the signal to noise ratio is significantly improved over  
20 the prior art proteins, such as wild-type GFP.

vii) Use as a cell or organelle integrity marker. By expressing the novel proteins targeted to an organelle, it is possible to calculate the leakage of the protein and use that as a measure of cell integrity.

25

viii) Use as a transfection marker, and as a marker to be used in combination with FACS sorting (eg. as described in Example 3). Due to the increased brightness of the novel proteins the quality of cell detection and sorting can be significantly improved.

30

ix) Use as real-time probe working at near physiological concentrations. Since the novel proteins of the present invention are significantly brighter than wtGFP when expressed in cells at about 37 °C and excited with light at about

-20-

490 nm, the concentration needed for visualization can be lowered. Target sites for enzymes engineered into the novel proteins, for example F64L-S175G-E222G-GFP, can therefore be present in the cell at low concentrations in living cells. This is important for two reasons: i) The probe must interfere as little as possible with the intracellular process being studied, and ii) the translational and transcriptional apparatus should be stressed minimally.

x) Transposon vector mutagenesis can be performed using the novel proteins as markers in transcriptional and translational fusions. Transposons may be used in microorganisms encoding the novel proteins. The transposons may be constructed for translational and transcriptional fusion to be used for screening for promoters. Transposon vectors encoding the novel proteins, for example F64L-S175G-E222G-GFP, can be used for tagging plasmids and chromosomes.

xi) Use as a reporter for bacterial detection by introducing the novel proteins into the genome of bacteriophages. By engineering the novel proteins, for example F64L-S175G-E222G-GFP, into the genome of a phage a diagnostic tool can be designed. F64L-S175G-E222G-GFP will be expressed only upon transfection of the genome into a living host. The host specificity is defined by the bacteriophage.

The invention is further illustrated by reference to the following examples and figures in which:

Figure 1 is the nucleotide Sequence of wtGFP (Chalfie et al, Science, (1994), 263, 802-5) and referred to herein as SEQ ID No.1.

Figure 2 is the corresponding amino acid sequence of wtGFP (Chalfie et al, Science, (1994), 263, 802-5) and referred to herein as SEQ ID No.2.

Figure 3 is the predicted amino acid sequence of F64L-V163A-E222G-GFP and referred to herein as SEQ ID No.3.

Figure 4 is the predicted amino acid sequence of F64L-S175G-E222G-GFP and referred to herein as SEQ ID No.4.

-21-

Figure 5 is the predicted amino acid sequence of F64L-S65T-S175G-GFP and referred to herein as SEQ ID No.5.

Figure 6 is the predicted amino acid sequence of F64L-S65T-V163A-GFP and referred to herein as SEQ ID No.6.

- 5 Figure 7 is a plot showing average fluorescence intensities of mutant GFPs according to the invention.

### EXAMPLES

10 1. Cloning of GFP gene and template vector construction

The GFP gene used in the present study was contained within the plasmid pGFP (Chalfie et al., *Science*, (1994), 263, 802-805; GenBank accession number U17997) obtained from Clontech Laboratories Inc. (Palo Alto, Ca, USA). The  
15 gene was amplified by PCR using Pfu polymerase (Promega, Madison, WI, USA) according to recognised protocols (Saiki et al., *Science*, (1988), 239, 487-491).

The sequences of primers used were:

GFP-1	5'-ggtacgggcccgcaccatgagtaaaggagaagaactttcac	SEQ ID No.7
GFP-2	5'-ggtacgggtaaccggttttgtatagttcatccatg	SEQ ID No.8
GFP-3	5'-ggtacgggcccgcaccatgggatccaaaggagaagaactttcac	SEQ ID No.9

- 20 Primer GFP-1 exhibits homology to the 5' region of the GFP gene and contains a partial Kozak site (Kozak, M, *Cell*, (1986), 44, 283) prior to the start codon for efficient initiation of translation in mammalian systems. Primer GFP-2 exhibits homology to the 3' region of the GFP gene and contains an additional  
25 Agel restriction enzyme site immediately prior to the stop codon to facilitate cloning of proteins by fusion to the C-terminus of GFP. Primer GFP-3 is similar to primer GFP-1 exhibiting homology to the 5' region of the GFP gene, but contains an additional restriction site (*Bam*HI) immediately after the initiation codon to facilitate cloning of proteins by fusion to the N-terminus of GFP. Amplified  
30 products resulting from PCR reactions containing primers GFP-1 and GFP-2, and GFP-3 and GFP-2 were tailed with a single 3'-deoxyadenosine using Taq

-22-

polymerase (Amersham Pharmacia Biotech, Amersham, UK) and ligated into the TA cloning vector pTARGET (Promega) according to manufacturer's instructions. The correct orientation relative to the CMV promoter and sequence of the insert was determined by automated DNA sequencing.

5

## 2. Generation of mutants of GFP

The following mutants of GFP were generated in the present study: F64L-GFP, V163A-GFP, S175G-GFP, E222G-GFP, F64L-E222G-GFP, F64L-V163A-GFP, F64L-S175G-GFP, V163A-S175G-GFP, V163A-E222G-GFP, S175G-E222G-GFP, F64L-S175G-E222G-GFP, V163A-S175G-E222G-GFP, F64L-V163A-E222G-GFP, F64L-S65T-S175G-GFP, F64L-S65T-V163A-GFP. Mutants of the GFP gene (SEQ ID 3) construct within pTARGET (See Example 1) were generated using the QuikChange™ site-directed mutagenesis kit (Stratagene, La Jolla, Ca, USA) according to manufacturer's instructions. The sequences of primers used to generate F64L, S65T, V163A, S175G and E222G single mutants have been documented in Table 1. Multiply-mutated GFP molecules were generated through successive mutagenesis reactions. All GFP mutant sequences were verified by automated sequencing.

20

Table 1

Primer	Mutation	Sequence (5' - 3')	SEQ ID No.
GFP-64f	F64L	ccaacactgtcactactctctttaggtgtcaat	10
GFP-64r	F64L	attgaacaccataagagagagtagtgacaagtgttg	11
GFP-65f	S65T	ccaacactgtcactactctcaactatggtgtcaatgctttca	12
GFP-65r	S65T	tgaaaagcattgaacaccataggtagagtagtgacaagtgttg	13
GFP-163f	V163A	gacaaacaaaagaatggaatcaaagccaactcaaaattagacac	14
GFP-163r	V163A	gtgtctaatttgaagttggcttgattccattctttgttgtc	15
GFP-175f	S175G	caacattgaagatggaggcggtcaactagcagacc	16
GFP-175r	S175G	ggctctgtagttgaacgcctccatcttcaatgttg	17
GFP-222f	E222G	ccacatggctctcttggcttgaacagctgctgg	18
GFP-222r	E222G	ccagcagctgttacaaagccaagaaggaccatgttg	19

-23-

3. Influence of individual mutations and combinations of F64L, S65T, V163A, S175G and E222G mutations upon GFP when expressed in mammalian cells

5 Plasmid DNA to be used for transfection was prepared for all GFP and EGFP constructs using the HiSpeed plasmid purification kit (Qiagen, Westberg, NL). DNA was diluted to 100 ng.  $\mu$ l<sup>-1</sup> in 18-Megohm water (Sigma) and 1  $\mu$ g used for transfections. For 50-80% confluency on the day of transfection, HeLa cells were plated at a density of  $5 \times 10^4$ /well in 6-well plates and incubated  
10 overnight. A 1:3 (1  $\mu$ g : 3  $\mu$ l) ratio of DNA to FuGene6 reagent (Roche) was used for each transient transfection reaction; 3  $\mu$ l FuGene6 was added to 87  $\mu$ l serum-free DMEM medium (Sigma) (containing penicillin/streptomycin, L-glutamine (GibcoBRL) and gently tapped to mix, then 10  $\mu$ l (1  $\mu$ g) construct DNA was added and again gently mixed. The FuGene6:DNA complex was incubated at room  
15 temperature for 40 minutes then added dropwise directly to the cells without changing the medium, and the plates swirled for even distribution.

Fluorescence measurements were made 24 or 48 hours after transfection. Briefly, the cells were washed in phosphate-buffered saline, released with the  
20 addition of 2 drops of Trypsin (GibcoBRL) and resuspended in 1 ml of complete DMEM medium (containing penicillin/streptomycin, L-glutamine and foetal bovine serum (Sigma). The cells were vortexed and analysed on a FACS Calibur flow cytometer (Becton Dickinson & Co., NJ, USA) for characterisation of whole cell fluorescence, with excitation at 488 nm and emission viewed with fluorescence  
25 filter set 530/30nm (range 515-545 nm). 10,000 events were collected for each transfection and 6-10 replicates carried out for each construct. Average fluorescent intensities from the FACS analysis were obtained as geometric means (mean fluorescence on log scale) and are shown in Figure 7.

30

Claims

1. A fluorescent protein which is derived from Green Fluorescent Protein (GFP) or any functional GFP analogue and has an amino acid sequence which is  
5 modified by amino acid substitution compared with the amino acid sequence of wild type Green Fluorescent Protein said modified fluorescent protein comprising:
  - i) an amino acid substitution at position F64;
  - ii) a single amino acid substitution at a position selected from the group consisting of positions S65 and E222; and
  - 10 iii) a single amino acid substitution at a position selected from the group consisting of positions V163 and S175;wherein said modified GFP has a different excitation spectrum and/or emission spectrum compared with wild type GFP.
- 15 2. A fluorescent protein according to claim 1 wherein the amino acid F at position 64 has been substituted by an amino acid selected from the group consisting of L, I, V, A and G.
- 20 3. A fluorescent protein according to claims 1 or 2 wherein the amino acid S at position 65 has been substituted by an amino acid selected from the group consisting of G, A, L, C, V, I and T.
4. A fluorescent protein according to claims 1 or 2 wherein the amino acid E at position 222 has been substituted by an amino acid selected from the group  
25 consisting of G, A, V, L, I, F, S, T, N and Q.
5. A fluorescent protein according to any of claims 1 to 4 wherein said V163 mutation is V163A.
- 30 6. A fluorescent protein according to any of claims 1 to 4 wherein said S175 mutation is S175G.

-25-

7. A fluorescent protein according to any of claims 1 to 6 selected from the group consisting of: F64L-V163A-E222G-GFP, F64L-S175G-E222G-GFP, F64L-S65T-S175G-GFP and F64L-S65T-V163A-GFP.
- 5 8. A fluorescent protein according to any of claims 1 to 7 having an amino acid sequence which is modified by amino acid substitution compared with the amino acid sequence of wild type Green Fluorescent Protein having the sequence: SEQ ID No.2.
- 10 9. A fluorescent protein derived from Green Fluorescent Protein (GFP) and having the amino acid sequence as set forth in SEQ ID No.3.
10. A fluorescent protein derived from Green Fluorescent Protein (GFP) and having the amino acid sequence as set forth in SEQ ID No.4.
- 15 11. A fluorescent protein derived from Green Fluorescent Protein (GFP) and having the amino acid sequence as set forth in SEQ ID No.5.
12. A fluorescent protein derived from Green Fluorescent Protein (GFP) and  
20 having the amino acid sequence as set forth in SEQ ID No.6.
13. A fusion compound comprising a protein of interest fused to a fluorescent protein said fluorescent protein being a modified protein according to any of claims 1 to 12.
- 25 14. A nucleic acid molecule comprising a nucleotide sequence encoding a fluorescent protein which is derived from Green Fluorescent Protein (GFP) or any functional GFP analogue and has an amino acid sequence which is modified by amino acid substitution compared with the amino acid sequence of wild type  
30 Green Fluorescent Protein said modified fluorescent protein comprising:
- i) an amino acid substitution at position F64;
  - ii) a single amino acid substitution at a position selected from the group consisting of positions S65 and E222; and

-26-

iii) a single amino acid substitution at a position selected from the group consisting of positions V163 and S175;

wherein said modified GFP has a different excitation spectrum and/or emission spectrum compared with wild type GFP.

5

15. A nucleic acid according to claim 14 encoding a fluorescent protein having an amino acid sequence which is modified by amino acid substitution compared with the amino acid sequence of wild type Green Fluorescent Protein having the sequence: SEQ ID No.2.

10

16. The nucleic acid molecule according to claims 14 or 15 encoding a fluorescent protein having an amino acid sequence selected from the group consisting SEQ ID No.3, SEQ ID No.4, SEQ ID No.5 and SEQ ID No.6.

15

17. A nucleic acid molecule comprising a nucleotide sequence encoding a fusion protein comprising a protein of interest fused to a fluorescent protein according to any one of claims 1 to 12.

20

18. An expression vector comprising suitable expression control sequences operably linked to a nucleic acid molecule according to any of claims 14 to 17.

19. A host cell transformed or transfected with a DNA construct comprising an expression vector according to claim 18.

25

20. The host cell according to claim 19 wherein said host cell is selected from the group consisting of a mammalian cell, a bacterial cell, a yeast cell and an insect cell.

30

21. A method for preparing a Green Fluorescent Protein (GFP) or a functional GFP analogue according to the present invention said method comprising cultivating a host according to claim 20 and obtaining therefrom the polypeptide expressed by said nucleotide sequence.



-27-

22. A method of measuring the expression of a protein of interest in a cell which method comprises:

- i) introducing into a cell a nucleic acid molecule comprising a nucleotide  
5 sequence encoding a fluorescent protein which is derived from Green  
Fluorescent Protein (GFP) or any functional GFP analogue according to any one  
of claims 1 to 12 said nucleic acid molecule being operably linked to and under  
the control of an expression control sequence which moderates expression of  
said protein of interest;
- 10 ii) culturing said cell under conditions suitable for the expression of said  
protein of interest; and
- iii) detecting the fluorescence emission of said Green Fluorescent Protein  
(GFP) or a functional GFP analogue as a means of measuring the expression of  
said protein of interest.

15

23. A method of determining the cellular and/or extracellular localisation of a protein of interest which method comprises:

- i) introducing into a cell a nucleic acid molecule comprising a nucleotide  
20 sequence encoding a fluorescent protein which is derived from Green  
Fluorescent Protein (GFP) or any functional GFP analogue according to any one  
of claims 1 to 12 fused to a nucleotide sequence encoding a protein of interest,  
said nucleic acid molecule being operably linked to and under the control of a  
suitable expression control sequence;
- 25 ii) culturing said cell under conditions suitable for the expression of said  
protein of interest; and
- iii) determining the cellular and/or extracellular localisation of said protein of  
interest by detecting the fluorescence emission by optical means.

30

24. A method of comparing the effect of one or more test substance(s) on the expression and/or localisation of one or more different protein(s) of interest in a cell which method comprises:

-28-

- i) introducing into a cell:
- a) a nucleic acid molecule comprising a nucleotide sequence encoding a Green Fluorescent Protein (GFP) or a functional GFP analogue according to any one of claims 1 to 12 optionally fused to a nucleotide sequence encoding a first protein of interest, said nucleic acid molecule being operably linked to and under the control of a first expression control sequence; and optionally,
- b) at least one different nucleic acid molecule encoding a protein reporter molecule optionally fused to a different protein of interest, each said nucleic acid molecule being operably linked to and under the control of a second expression control sequence wherein said protein reporter molecule has or is capable of generating an emission signal which is spectrally distinct from that of said Green Fluorescent Protein (GFP) or functional GFP analogue;
- ii) culturing said cells under conditions suitable for the expression of said protein(s) of interest in the presence and absence of said test substance(s);
- iii) determining the expression and/or localisation of said protein(s) of interest in said cells by detecting the fluorescence emission by optical means; and
- iv) comparing the fluorescence emission obtained in the presence and absence of said test substance(s) to determine the effect of said test substance(s) on the expression and/or localisation of said protein(s) of interest.

20

25. The method according to claim 24 wherein samples of said cells in a fluid medium are introduced into separate vessels for each of said test substances to be studied.

Figure 1/7Nucleotide Sequence of wtGFP (Chalfie et al. Science, (1994), 263, 802-5):SEQ ID No.1

```
atg agt aaa gga gaa gaa ctt ttc act gga gtt gtc cca att ctt gtt 48
gaa tta gat ggt gat gtt aat ggg cac aaa ttt tct gtc agt gga gag 96
ggt gaa ggt gat gca aca tac gga aaa ctt acc ctt aaa ttt att tgc 144
act act gga aaa cta cct gtt cca tgg cca aca ctt gtc act act ttc 192
tct tat ggt gtt caa tgc ttt tca aga tac cca gat cat atg aaa cgg 240
cat gac ttt ttc aag agt gcc atg ccc gaa ggt tat gta cag gaa aga 288
act ata ttt ttc aaa gat gac ggg aac tac aag aca cgt gct gaa gtc 336
aag ttt gaa ggt gat acc ctt gtt aat aga atc gag tta aaa ggt att 384
gat ttt aaa gaa gat gga aac att ctt gga cac aaa ttg gaa tac aac 432
tat aac tca cac aat gta tac atc atg gca gac aaa caa aag aat gga 480
atc aaa gtt aac ttc aaa att aga cac aac att gaa gat gga agc gtt 528
caa cta gca gac cat tat caa caa aat act cca att ggc gat ggc cct 576
gtc ctt tta cca gac aac cat tac ctg tcc aca caa tct gcc ctt tct 624
aaa gat ccc aac gaa aag aga gac cac atg gtc ctt ctt gag ttt gta 672
aca gct gct ggg att aca cat ggc atg gat gaa cta tac aaa tag 717
```



Figure 2/7Amino Acid Sequence of wtGFP (Chalfie et al. Science, (1994), 263, 802-5SEQ ID No.2

```

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
1          5          10          15
Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
20          25          30
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
35          40          45
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
50          55          60
Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Arg
65          70          75          80
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
85          90          95
Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
100          105          110
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
115          120          125
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn
130          135          140
Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly
145          150          155          160
Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
165          170          175
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
180          185          190
Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
195          200          205
Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val
210          215          220
Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
225          230          235

```



Figure 3/7Predicted Amino Acid Sequence of F64L-V163A-E222G-GFP:SEQ ID No.3

```

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
1          5          10          15
Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
20          25          30
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
35          40          45
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu
50          55          60
Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Arg
65          70          75          80
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
85          90          95
Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
100         105         110
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
115         120         125
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn
130         135         140
Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly
145         150         155         160
Ile Lys Ala Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
165         170         175
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
180         185         190
Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
195         200         205
Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Gly Phe Val
210         215         220
Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
225         230         235

```





Figure 4/7Predicted Amino Acid Sequence of F64L-S175G-E222G-GFP:SEQ ID No.4

```

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
1           5           10           15
Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
20           25           30
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
35           40           45
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu
50           55           60
Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Arg
65           70           75           80
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
85           90           95
Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
100          105          110
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
115          120          125
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn
130          135          140
Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly
145          150          155          160
Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Gly Val
165          170          175
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
180          185          190
Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
195          200          205
Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Gly Phe Val
210          215          220
Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
225          230          235

```



Figure 5/7Predicted Amino Acid Sequence of F64L-S65T-S175G-GFP:SEQ ID No.5

```

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
1           5           10           15
Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
20           25           30
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
35           40           45
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu
50           55           60
Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Arg
65           70           75           80
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
85           90           95
Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
100          105          110
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
115          120          125
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn
130          135          140
Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly
145          150          155          160
Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Gly Val
165          170          175
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
180          185          190
Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
195          200          205
Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val
210          215          220
Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
225          230          235

```



Figure 6/7Predicted Amino Acid Sequence of F64L-S65T-V163A-GFP:SEQ ID No.6

```

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
1           5           10           15
Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
20           25           30
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
35           40           45
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu
50           55           60
Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Arg
65           70           75           80
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
85           90           95
Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
100          105          110
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
115          120          125
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn
130          135          140
Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly
145          150          155          160
Ile Lys Ala Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
165          170          175
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
180          185          190
Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
195          200          205
Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val
210          215          220
Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
225          230          235

```



Figure 7/7